LOCAL MASS DISTRIBUTION PARTITIONING FOR OBJECT RECOGNITION

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001]   This is the first application filed for the present invention.

FIELD OF THE INVENTION

[0002]   The present invention relates to machine vision, and in particular to a global method for object recognition using local moments.

BACKGROUND OF THE INVENTION

[0003]   Techniques of optical object recognition have applications in a variety of fields, including automated manufacturing, biomedical engineering, security, and document analysis.

[0004]   Most object recognition methods measure the similarity between two objects based on a set of features extracted from each of the objects. These methods can be classified into three types—*global feature methods, structural feature methods,* and *relational graph methods*—as explained in "Model-Based Recognition in Robot Vision" (R.T. Chin and C.R. Dyer, Computing Surveys, Vol.18, No.1, March 1996).

[0005]   In global feature methods, the object is represented by a *feature vector* containing n global features of the object. Thus, each object can be viewed as a point in an n-dimensional feature space, and two objects

- 2 -

can be compared using a function of their corresponding points in the feature space, such as a measure of distance.

[0006] An application usually associated with global methods is object identification, which consists of recognizing an isolated object (target object) contained in an image as one in a bank of predefined model objects. Typically, the object identification process comprises two main phases: model definition and object recognition. During the model definition phase, each of the model objects is represented by a feature vector according to a chosen feature extraction scheme. During the object recognition phase, an image is acquired (e.g. using a gray-scale digital camera, etc.) and pre-processed to isolate individual objects. For one of these isolated objects, the target object, a feature vector is extracted according to the chosen feature extraction scheme. This feature vector is then compared to the feature vectors of the model objects; each target-model pair is attributed a similarity score.

[0007] In object identification applications, the model object as it occurs in the target image may have undergone an affine transformation (e.g. translation, rotation, scale variation, aspect-ratio variation, skew, etc.) and degradation of various types (e.g. erosion and dilatation, addition or subtraction of pieces, etc.). For gray scale images, the model object may have been affected by a change in image contrast or mean gray level.

[0008] Therefore, object identification must support the set of affine transformations and be robust to the types of degradation expected for the specific application. Support for the transformations can be achieved by including in the pre-processing stage one or

more standardization steps during which the object is roughly centered in the sub-image and normalized to a standard size and orientation; and/or by selecting features that are invariant under the transformations; and/or by choosing a similarity measure that detects the correlation between the vectors of two objects related through an affine transformation.

[0009]   In global feature methods, the process of feature extraction consists of extracting from the object a set of features adequate for identification purposes, namely a set of features that vary across the model objects and thus enable their differentiation. Various types of features are used in object recognition, as described in "Feature Extraction Methods for Character Recognition—A Survey" (Ø.D. Trier, A.K. Jain and T. Taxt, Pattern Recognition, Vol. 29, No. 4, pp. 641-662, 1996).

[0010]   A particular class of features used for object recognition are those derived from geometric moments of the object, as described in *Moment Functions in Image Analysis* (R. Mukundan and K.R. Ramakrishnan, World Scientific Publishing, 1998). Geometric moments of different orders of the object provide different spatial characteristics of the intensity distribution within the object (and of the mass distribution of the object for binary images); for example, moments of order zero and one together provide elementary object descriptors such as the total intensity and the intensity centroid of the object (total mass and center of mass for binary images).

[0011]   Geometric moments have many desirable characteristics for use as features in object recognition. *Central moments* that are calculated by shifting the origin of the reference system to the intensity centroid are

invariant to translation. Furthermore, geometric moments can be combined to obtain moment features that are invariant to other transformations such as uniform scale variation, aspect-ratio variation and rotation, called geometric moment invariants.

[0012]   In typical moment-based methods of object recognition, an object is represented by a set of features derived from global moments of the object of various orders. Increasing the order of the moments used provides a larger set of features to represent the object, but also results in a greater sensitivity to noise.

[0013]   An alternative feature set that does not require the use of higher order moments uses local moments of the object, as opposed to global moments. In a method known as "Zoning" applied to binary images, an $n \times m$ (uniform rectangular) grid is superimposed on the object image, and the masses (zero-order moments) of the object in each of the $n \times m$ regions are used as features. However, as the outer boundary of the grid is fixed by the outmost points of the object, the addition or subtraction of a piece, even of negligible mass relative to the object, can significantly alter the grid and therefore the extracted features.

[0014]   Another method based on local moments is described in "Scene Classification by Fuzzy Local Moments" (H.D. Cheng and R. Desai, International Journal of Pattern Recognition and Artificial Intelligence, Vol. 12, No. 7, 921-938,1998). The method locates the intensity centroid of the object, on which it centers an $n \times m$ radial grid; the outer boundary of the grid, the nth concentric circle, is fixed by the object point at the greatest radial distance from the center of mass. The features are the

- 5 -

first-order radial moments calculated in each of the $n \times m$ regions. This method suffers from the same disadvantage as zoning, namely that the size and position of the grid is fixed by the outer boundary of the object, which is vulnerable to degradation. Also, using a radial system of axes is not computationally efficient.

[0015]   There is therefore a need in the industry for an object recognition method that will overcome the above-identified drawbacks.

SUMMARY OF THE INVENTION

[0016]   Accordingly, it is an object of the invention to provide a method used in object recognition that offers the benefits of local moment-based techniques but with an improved robustness to noise.

[0017]   According to an embodiment of the invention, there is provided a method for determining a similarity score of a target object with respect to a model object. The target object is in a plane and the model object is represented by a model feature vector. The method comprises generating regions of the plane according to a first mass distribution of the target object and a second mass distribution of a part of the target object. Each of the generated regions has a corresponding mass distribution indicator. The method further comprises calculating a target feature vector for the target object according to at least one of the corresponding mass distribution indicators. Finally, the method computes the similarity score using the target feature vector and the model feature vector.

- 6 -

BRIEF DESCRIPTION OF THE DRAWINGS

[0018]   The detailed description of the invention will refer to the appended drawings, in which:

[0019]   FIG.1 is a schematic illustrating an exemplary application of the present invention, namely object identification;

[0020]   FIG. 2 is a flow chart illustrating principal steps of the model definition phase in accordance with the present invention;

[0021]   FIG. 3 is a flow chart illustrating principal steps of the process of generating regions, whether during the model definition or the object recognition phase, in accordance with an embodiment of the present invention;

[0022]   FIG. 4 is a flow chart illustrating principal steps of the process of partitioning the plane containing the object, whether during the model definition or the object recognition phase, in accordance with an embodiment of the present invention;

[0023]   FIG. 5a-d is a schematic sequentially illustrating successive steps in the process of generating regions in accordance with FIG. 4 for the model definition phase;

[0024]   FIG. 6 is a flow chart illustrating principal steps in the process of extracting a feature vector, whether during the model definition or the object recognition phase, in accordance with an embodiment of the present invention;

- 7 -

[0025]   FIG. 7 is a schematic illustrating the process of FIG. 6 for the model definition phase;

[0026]   FIG. 8 is a flow chart illustrating principal steps of the object recognition phase in accordance with the present invention;

[0027]   FIG. 9a-b is a schematic illustrating the processes of generating regions and extracting a feature vector in accordance with FIGS. 4 and 6 for the object recognition phase; and

[0028]   FIG. 10 is a flow chart illustrating principal steps in the process of computing the similarity score in accordance with an embodiment of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

[0029]   FIG. 1 schematically illustrates an exemplary application of the present invention, namely object identification. Object identification is used for Optical Character Recognition or for the recognition of mechanical parts and tools in automated manufacturing, for example. Referring to FIG. 1, object identification consists of identifying an isolated object 21, the target object, in an image 20 as one in a bank of predefined model objects 22. In this application, a similarity score is computed for each of the target-model object pairs. These scores can then be used to select, among the model objects 22, one or more candidates for the target object 21.

[0030]   The method of the present invention comprises two main phases, namely model definition and object recognition. The model definition phase involves

- 8 -

representing the model object by a feature vector, according to a chosen feature extraction scheme. The object recognition phase involves extracting from a target object contained in an image a feature vector according to the chosen feature extraction scheme, and computing a similarity score between this feature vector and the feature vector of the model object.

[0031]   Model Definition

[0032]   In the present invention, the model definition phase (FIG. 2) comprises two successive steps: generation of regions of a plane containing a model object (at 200) and extraction of features from these regions (at 202).

[0033]   The model object can be a suitable image of the object (i.e. an image in which the physical object is neither degraded nor occluded by other objects and taken under suitable conditions of lighting, contrast, noise level, etc.) acquired by an image acquisition device (e.g. gray scale digital camera, electromagnetic or ultra-sonic imaging system, etc.). The acquired image is often converted to another representation, such as a binary representation, or an edge, skeleton or crest representation. Alternatively, the model object can take the form of a so-called "synthetic" description of the object, such as a drawing produced using a software program or a set of mathematical equations describing the mass distribution (or intensity distribution) of the object. In any case, the model object is defined as a set of weighted points in a plane with an associated plane coordinate system. FIG. 5a shows an exemplary model object 23, which for the purpose of the example, is a set of points of equal

- 9 -

weight. The model object is in a plane 24 with an associated plane coordinate system 25.

[0034]   Referring to FIG. 2, the model definition phase begins with generating regions of the plane containing the model object (at 200). As shown in FIG. 3, regions are generated by first partitioning the plane (using two successive partitions at 300 and 302) and then, selecting regions based on this partition (at 304). The first partition (at 300) partitions the plane according to a first mass distribution of the object thereby producing global parts.   The second partition (at 302) partitions these global parts according to a second mass distribution of a part of the object thereby producing disjoint parts. A region generated from the total partition is defined as any set of one or more of these disjoint parts, and the process of selecting regions based on the partition (at 304) consists of selecting any set of such regions.

[0035]   The partitioning of the plane for the preferred embodiment is described with reference to FIGS. 3, 4 and 5. Referring to FIG. 5b, for the first partition (at 300), the method locates the axes 27 parallel to the axes of the plane coordinate system 25 passing through the center of mass 26 of the object 23 (at 400). Referring to FIG. 5c, for the second partition (at 302), the method selects each of the four half-planes bordered by a center of mass axis 27 (at 402). An exemplary half-plane 28 is illustrated in FIG. 5c. For each half-plane, as illustrated for the half-plane 28, the method locates the axis 30 parallel to the bordering center of mass axis 27 passing through the center of mass 29 of the half-plane (at 404); this local center of mass axis is used to partition the quadrants that form the half-plane. As shown in FIG. 5d,

- 10 -

the set of global center of mass axes 27 and local center of mass axes 30 define a total partition of the plane into disjoint parts(at 408), and the set of regions selected consists of all the regions 31 obtained by combining two adjacent disjoint parts.

[0036]   In another preferred embodiment of the invention, the second partition (at 302) is performed by selecting each of the quadrants bordered by the center of mass axes, and for each of these quadrants, locating the axes parallel to the center of mass axes passing through the center of mass of the quadrant. Again, the global and local center of mass axes define a partition of the plane into disjoint parts. In this embodiment, the set of regions selected consists of all the disjoint parts resulting from the partition.

[0037]   Other indicators of the distribution of mass of the object may be used to partition the object, such as other indicators derived from geometric moments, or alternatively the median.

[0038]   Referring to FIG. 2, the second step of the model definition phase consists of creating a feature vector for the model object (at 202), by extracting features from the generated regions. An extracted feature can be any indicator of the distribution of mass of the object in the corresponding region, such as an indicator derived from geometric moments of various orders, or alternatively from the median.

[0039]   In a preferred embodiment, each feature extracted from a region is a coordinate (x or y) of the center of mass of the object in this region. Extraction of the feature vector for the preferred embodiment is

- 11 -

described with reference to FIGS. 6 and 7. Referring to FIG. 7, for each of the n generated regions 31, the center of mass 32 $\overline{M}(\overline{x}_i^M, \overline{y}_i^M)$ of the model object in this region is calculated (at 602). The feature vector V^M is then created using, for feature vector elements, the coordinates of the local centers of mass 32 of the object in the n generated regions represented in a coordinate system located at the global center of mass 26 (at 606).

$$V^M = \left(\overline{x}_1^M, \overline{y}_1^M, \overline{x}_2^M, \overline{y}_2^M, \ldots \overline{x}_n^M, \overline{y}_n^M\right)$$

[0040]   Object Recognition

[0041]   During the object recognition phase, a feature vector is extracted from a target object contained in an image according to the same feature extraction scheme as used during the model definition phase, and a similarity score is then computed between this feature vector and the feature vector of the model object.

[0042]   Preliminary steps of the object recognition phase include acquiring and pre-processing an image to obtain the target object.' First, an image of a scene including a target object for comparison with a model object is acquired using an image acquisition device, such as any of those listed above for obtaining the model object. The model object as it occurs in the target image may have undergone certain affine transformations (e.g. translation, rotation, scale variation, aspect-ratio variation) and degradation of various types (e.g. erosion and dilatation, addition or subtraction of pieces, etc.). Second, the image is pre-processed. Objects in the image are isolated using any method known in the art (e.g. blob

- 12 -

analysis, etc.). One or more standardization steps may be applied to the image for the method to support some of the expected transformations; for example, the image may be scaled or rotated for the image objects to appear in a standard size and orientation. Also, the image may be converted to another representation, such as a binary representation, or an edge, skeleton or crest representation. In any case, each isolated object is defined, like the model object, as a set of weighted points; these isolated objects appear in a plane with an associated coordinate system. In a final pre-processing step, a target object is selected among the isolated objects.

[0043]   FIG. 9a illustrates an exemplary target object 33 and another isolated object 36 in a plane 34 with an associated coordinate system 35. For the purpose of this example, the isolated objects consist of sets of points of equal weight. The target object 33 is an occurrence of the model object 23 that has undergone translation, a variation in scale, and degradation, in particular the addition of a long piece 37.

[0044]   FIG. 8 is a flow chart illustrating the principal steps of the object recognition phase in accordance with the present invention. The feature vector is extracted from the target object according to the same feature extraction scheme as used for the model definition phase. Thus, referring to FIG. 8, regions of the plane containing the target object are generated (at 800) and then features are extracted from these regions (at 802), following the corresponding steps of the model definition phase (at 200 and 202). The final step consists of computing the similarity score between the target object

10442-32US

- 13 -

feature vector and the model object feature vector (at 804).

[0045]   Regions are generated (at 800) in exactly the same way as during the model definition phase (at 200). Referring to FIG. 3, the plane containing the target object is partitioned (using two successive partitions at 300 and 302), and a set of regions based on this partition is selected (at 304). In a preferred embodiment, regions of the plane are generated as described in FIG. 4. Referring to FIG. 9b, the set of global center of mass axes 39 and local center of mass axes 40 define a partition of the plane 34 containing the target object 33. The set of generated regions consists of all the regions 41 obtained by combining two adjacent disjoint parts of the partition.

[0046]   The feature vector is extracted from the generated regions (at 802) in precisely the same way as during the model definition phase (at 202). Therefore, in a preferred embodiment, the feature vector is created as described in FIG. 6. Referring to FIG. 9b, for each of the n generated regions 41, the center of mass 42 $\overline{M}(\overline{x}_i^T, \overline{y}_i^T)$ of the target object in this region is calculated (at 602). The feature vector V^T is then created using, for feature vector elements, the coordinates of the local centers of mass of the object in the n generated regions represented in a coordinate system located at the global center of mass 38 (at 606).

$$V^T = \left(\overline{x}_1^T, \overline{y}_1^T, \overline{x}_2^T, \overline{y}_2^T, \dots \overline{x}_n^T, \overline{y}_n^T \right)$$

[0047]   In the proposed feature extraction scheme, both the method of partitioning and the features themselves present advantages.

[0048]   First, the method of partitioning a plane containing the object using successive partitions based on local mass distributions of the object is more robust to degradation of the boundary of the object than other methods known in the art. In the present method, a piece added to or subtracted from the object affects the partition proportionally to its mass relative to the object. Thus, a piece of relatively small mass (e.g. 37) will not significantly affect the overall partition or the extracted features as a whole; its effect will be largely restricted to the regions containing the addition or subtraction of mass. Second, the extracted features are derived from first and zero order moments, which are less sensitive to noise than moments of higher order. The local moments are represented in a coordinate system attached to the object and therefore are invariant to translation of the object in the image. Also, using orthogonal coordinate axes parallel to the natural axes of the image is more computationally effective than using a radial system of axes, for example.

[0049]   The final step of the object recognition phase consists of computing a similarity score between the target and model object feature vectors (at 804). This similarity score is computed using a set of weights assigned to the features. In a preferred embodiment of the invention, all the weights are equal to one, which is equivalent to not using weights.

[0050]   A commonly used measure of the correspondence between two moment-based feature vectors is

Euclidean distance. The Euclidean distance D between the target and model object feature vectors V^T and V^M is given by:

$$D(V^T, V^M) = \left| V^T - V^M \right| = \sqrt{\sum_{i=1}^{n} \left[(\bar{x}_i^T - \bar{x}_i^M)^2 + (\bar{y}_i^T - \bar{y}_i^M)^2 \right]}$$

[0051]   This measure is simple to compute, but for the method to support a scale difference between the target and model objects, one of the objects or feature vectors must be scaled beforehand.

[0052]   An alternative measure that is scale invariant is related to the angle θ between the target and model object feature vectors. The scale invariant similarity score S is given by:

$$S(V^T, V^M) = \cos^2(\theta) = \left(\frac{V^T \cdot V^M}{\left| V^T \right| \left| V^M \right|} \right)^2$$

[0053]   However, this score does not support independent scale variations in the x and y directions or, equivalently, transformations affecting the aspect ratio of the object.

[0054]   FIG. 10 illustrates the principal steps of the process of computing an aspect ratio invariant score, in accordance with the preferred embodiment of the invention. First (at 1000), the target and model objects are each represented by two independent feature vectors, the first containing only the x coordinates of the local

- 16 -

centers of mass and the other containing only the y coordinates:

$$V_x^T = (\bar{x}_1^T, \bar{x}_2^T, ..\bar{x}_n^T) \qquad V_y^T = (\bar{y}_1^T, \bar{y}_2^T, ..\bar{y}_n^T)$$

$$V_x^M = (\bar{x}_1^M, \bar{x}_2^M, ..\bar{x}_n^M) \qquad V_y^M = (\bar{y}_1^M, \bar{y}_2^M, ..\bar{y}_n^M)$$

[0055]   Second (at 1002), two independent similarity scores S_x and S_y are computed for the x and y feature vectors respectively:

$$S_x = S(V_x^T, V_x^M) = \cos^2(\theta_x) = \left(\frac{V_x^T \cdot V_x^M}{\left|V_x^T\right|\left|V_x^M\right|}\right)^2$$

$$S_y = S(V_y^T, V_y^M) = \cos^2(\theta_y) = \left(\frac{V_y^T \cdot V_y^M}{\left|V_y^T\right|\left|V_y^M\right|}\right)^2$$

[0056]   Finally (at 1004), the global similarity score between the target and model objects is computed as the average of the two independent similarity scores:

$$S(V^T, V^M) = \frac{S_x + S_y}{2}$$

[0057]   In another embodiment of the invention, non-trivial weights may be used. During a preliminary stage, each pair of corresponding features $\bar{x}_i^T$ and $\bar{x}_i^M$ ($\bar{y}_i^T$ and $\bar{y}_i^M$) of the target and model objects is assigned a weight w_{ix} (w_{iy}).

- 17 -

$$W = \begin{bmatrix} w_{1x} & 0 & 0 & 0 & & 0 \\ 0 & w_{1y} & 0 & 0 & & 0 \\ 0 & 0 & w_{2x} & 0 & & 0 \\ 0 & 0 & 0 & w_{2y} & & 0 \\ & & & & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & w_{ny} \end{bmatrix}$$

[0058]   The weight matrix can be used with any of the similarity scores described previously. The weighted Euclidean distance D_W and the weighted scale invariant score S_W are given by:

$$D_W(V^T, V^M) = \left| V^T - V^M \right|_W = \sqrt{\sum_{i=1}^{n} w_i \left[(\bar{x}_i^T - \bar{x}_i^M)^2 + (\bar{y}_i^T - \bar{y}_i^M)^2 \right]}$$

$$S_W(V^T, V^M) = \left(\frac{(V^T \cdot V^M)_W}{|V^T|_W |V^M|_W} \right)^2$$

where

$$|A|_W = \sqrt{A'WA}$$

$$(A \cdot B)_W = A'WB$$

where A' denotes the transpose of A.

[0059]   Similarly, a weighted aspect ratio invariant score can be obtained by averaging the weighted similarity scores S_{xW_x} and S_{yW_y} for the x and y feature vectors given by:

- 18 -

$$S_{x\overline{W}_x}(V^T, V^M) = \left(\frac{\left(V_x^T \cdot V_x^M\right)_{\overline{W}_x}}{\left|V_x^T\right|_{\overline{W}_x}\left|V_x^M\right|_{\overline{W}_x}}\right)^2 \quad \text{where} \quad W_x = \begin{bmatrix} w_{1x} & 0 & 0 & 0 \\ 0 & w_{2x} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & w_{nx} \end{bmatrix}$$

$$S_{y\overline{W}_y}(V^T, V^M) = \left(\frac{\left(V_y^T \cdot V_y^M\right)_{\overline{W}_y}}{\left|V_y^T\right|_{\overline{W}_y}\left|V_y^M\right|_{\overline{W}_y}}\right)^2 \quad \text{where} \quad W_y = \begin{bmatrix} w_{1y} & 0 & 0 & 0 \\ 0 & w_{2y} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & w_{ny} \end{bmatrix}$$

[0060]   An application of the present invention is object identification, which consists of matching a target object contained in an image to one or more model objects contained in a bank of model objects. In this case, similarity scores are computed between the target object and each of the model objects, and these scores are then used to select one or more model objects based on certain criteria.

[0061]   Given a specific bank of model objects, some of the extracted features vary more over the set of model objects than others and therefore contribute more to differentiating the similarity scores computed between a target object and each of the model objects. In order to accentuate the differences between the similarity scores, larger weights can be assigned to features associated with greater model object variability. In a preferred embodiment, a common weight matrix is assigned to the bank of model objects during a training stage.

[0062]   The embodiments of the invention described above are intended to be exemplary only. The scope of the invention is therefore intended to be limited solely by the scope of the appended claims.